Proposal to add Arabic letter for Swahili

Denis Moyogo Jacquerye <moyogo@gmail.com>
2024-02-11

This document requests the addition of an Arabic letter character (ARABIC LETTER YEH WITH FOUR DOTS BELOW) used in 20th century manuscripts of Muhamadi Kijuma in Swahili (ISO 639-3: swa). The letter was used for the palatal nasal, ⟨ny⟩ in Latin orthography.

Muhammad bin Abubakar bin Umar al-Bakriy (1885-1845), known as Muhamadi Kijuma (or Kijumwa), was a writer, poet, translator in the Swahili language. His manuscripts are collected in Kenya, Tanzania and in Europe. In 2022, professor of Literatures African Languages at University of Bayreuth, Clarissa Vierke, co-author of Muhamadi Kijuma – Texts from the Dammann Papers and other Collections 2010, said "[Muhamadi Kijuma] supplied many European researchers with manuscripts of Swahili poetry and introduced also them to the manuscript culture and Islamic poetry. He truly shaped our understanding of Swahili in Arabic script and most of the manuscripts stored in London, Hamburg, or Berlin nowadays are in his handwriting, including glossaries and explanations."

Character to be added

<table>
<thead>
<tr>
<th>Codepoint</th>
<th>Isolated</th>
<th>Final</th>
<th>Medial</th>
<th>Initial</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>U+10EC7</td>
<td>ݑ</td>
<td>ݑ</td>
<td>ݙ</td>
<td>ݙ</td>
<td>ARABIC LETTER YEH WITH FOUR DOTS BELOW</td>
</tr>
</tbody>
</table>

Annotations and properties

NamesLists.txt

The following annotations are recommended for NamesList.txt:

10EC7 ARABIC LETTER YEH WITH FOUR DOTS BELOW
   * Swahili manuscripts

UnicodeData.txt

10EC7;ARABIC LETTER YEH WITH FOUR DOTS BELOW;Lo;0;AL;;;;N;;;;;

ArabicShaping.txt

10EC7; DOTLESS YEH WITH 4 DOTS BELOW; D; YEH
Confusables

The ARABIC LETTER YEH WITH FOUR DOTS BELOW has forms confusable with forms of ٥ U+0680 ARABIC LETTER BEHEH.

Examples

ny palatal /n/ nyama ‘meat’
t cerebral [retroflex] /t/ t̪̬oto ‘small’ [= /ch/ Kisiu, marked by Dammann, but written in the texts only as /t/ without diacritic]
v dentilabial voiced /v/ viva ‘ripen’ (Kiunguja: ɪvɑ)
c chachache ‘few’ (Kiunguja: chache)
mb k̪ambɑ ‘crab’
ranji ranji ‘colour’. M.K. probably also pronounced the adjective -ema in class 5 ‘good’ as jema
nd with cerebral pronunciation ndimo ‘it is inside’. In Lamu one most commonly hears /nd/.
d with interdental or postdental pronunciation. The only example given by M.K. is the nasal compound in m(ɑ)pɛndi ‘lover’.
p pipa ‘barrel’

A note on Arabic script: SBB Hs. or. 9910

Figure 1. ARABIC LETTER YEH WITH FOUR DOTS BELOW and its transliteration ⟨ny⟩ in Miehe & Vierke 2010, p. 169 in “ny” and “nyama”.
Figure 2. ARABIC LETTER YEH WITH FOUR DOTS BELOW in Qiṣṣatī Yūsufū (Hs. or: 9893), p. 1.
Figure 3. ARABIC LETTER YEH WITH FOUR DOTS BELOW in Qiṣṣati Yūsufu (Hs. or. 9893), p. 2.
Figure 4. ARABIC LETTER YEH WITH FOUR DOTS BELOW in Qissat Qudhi (Ms. or 9893), p. 73. Note: the final form has a complete shape in the next example from p. 76.
Figure 5. ARABIC LETTER YEH WITH FOUR DOTS BELOW in Qīṣṣat Qādhi (Hs. or. 9893), p. 76.
Figure 6. ARABIC LETTER YEH WITH FOUR DOTS BELOW in Utendi wa Safāri (SBB Hs. or. 9896) in “Tukafanya” and “kuonyana”. Note: the final form has a complete shape in the next example from the following page.
Figure 7. ARABIC LETTER YEH WITH FOUR DOTS BELOW in *Utendi wa Safāri* (SBB Hs. or. 9896) in “kinyezi” and “Mwenye”.
Figure 8. ARABIC LETTER YEH WITH FOUR DOTS BELOW in Swahili Notes and Glosses (MS 47796a) in “nyoka”, “nyende” [nyenje], “nyani”, etc.
References by Muhamadi Kijuma

- *Qiṣṣati Yūsufu* (Hs. or. 9893). 1913. http://resolver.staatsbibliothek-berlin.de/SBB00015E2700000000
- *Qiṣṣat Qādhi* (Hs. or. 9893). 1937. http://resolver.staatsbibliothek-berlin.de/SBB00015E2700000000
- *Swahili Notes and Glosses (MS 47796a)*. 1935.
  http://digital.soas.ac.uk/LSMD000192/00001
- *Utendi wa Safāri* (SBB Hs. or. 9896). 1937.

Other references

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from
http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html for guidelines and details before filling this form.
Please ensure you are using the latest Form from
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest Roadmaps.

A. Administrative

1. Title: Proposal to add Arabic letter for Swahili
2. Requester's name: Denis Moyogo Jacquerye <moyogo@gmail.com>
3. Requester type (Member body/Liaison/Individual contribution): Individual contribution
4. Submission date: 2024-02-11
5. Requester's reference (if applicable):
6. Choose one of the following:
   a. This is a complete proposal: YES
   b. More information will be provided later: NO

B. Technical - General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters):
   b. The proposal is for addition of character(s) to an existing block:

2. Number of characters in proposal: 1
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary X B.1-Specialized (small collection)
   B.2-Specialized (large collection) C-Major extinct
   D-Attested extinct E-Minor extinct
   F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? YES
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
   Denis Moyogo Jacquerye <moyogo@gmail.com>
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? YES
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? YES

7. Special encoding issue
   Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? YES (presentation, sorting)

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see UAX#44: http://www.unicode.org/reports/tr44/ and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Has this proposal for addition of character(s) been submitted before?</td>
<td>NO</td>
</tr>
<tr>
<td>2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?</td>
<td>YES</td>
</tr>
<tr>
<td>3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?</td>
<td>YES</td>
</tr>
<tr>
<td>4. The context of use for the proposed characters type of use; common or rare)</td>
<td>YES</td>
</tr>
<tr>
<td>5. Are the proposed characters in current use by the user community?</td>
<td>YES</td>
</tr>
<tr>
<td>6. After giving due considerations to the principles in the P&amp;P document must the proposed characters be entirely in the BMP?</td>
<td>YES</td>
</tr>
<tr>
<td>7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?</td>
<td>YES</td>
</tr>
<tr>
<td>8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?</td>
<td>NO</td>
</tr>
<tr>
<td>9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?</td>
<td>NO</td>
</tr>
<tr>
<td>10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?</td>
<td>NO</td>
</tr>
<tr>
<td>11. Does the proposal include use of combining characters and/or use of composite sequences?</td>
<td>NO</td>
</tr>
<tr>
<td>12. Does the proposal contain characters with any special properties such as control function or similar semantics?</td>
<td>NO</td>
</tr>
<tr>
<td>13. Does the proposal contain any Ideographic compatibility characters?</td>
<td>NO</td>
</tr>
</tbody>
</table>